Agency, simulation and self-identification — Source link

Marc Jeannerod, Elisabeth Pacherie

Institutions: Centre national de la recherche scientifique

Related papers:

- Philosophical conceptions of the self: implications for cognitive science

- Beyond consciousness of external reality: A "who" system for consciousness of action and self-consciousness

- How the body shapes the mind

- Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis.

- Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients

Agency, Simulation and Self-identification

Marc Jeannerod, Elisabeth Pacherie

Agency, Simulation and Self-identification

MARC JEANNEROD AND ELISABETH PACHERIE

Abstract: This paper is concerned with the problem of self-identification in the domain of action. We claim that this problem can arise not just for the self as object, but also for the self as subject in the ascription of agency. We discuss and evaluate some proposals concerning the mechanisms involved in self-identification and in agency-ascription, and their possible impairments in pathological cases. We argue in favor of a simulation hypothesis that claims that actions, whether overt or covert, are centrally simulated by the neural network, and that this simulation provides the basis for action recognition and attribution.

In this paper we will be concerned with the problem of self-identification as it arises in the domain of action, which we will take to include both overt and covert or simulated actions.

Talk of a problem of identification presupposes a contrast set, and the possibility that, in seeking to identify one self, one picks out something in the contrast set instead. With self-identification, two contrast sets must be considered: the world at large and the set of other selves. The problem of self-identification therefore divides into two sub-problems: self-world discrimination and self-other discrimination. We focus here on the latter problem. Our first aim is to show that there is a problem of self-other discrimination. Our second aim is to discuss and evaluate some proposals concerning the mechanisms involved in self-identification and in agency-ascription, and their possible impairments in pathological cases.

In the first section, we will examine some of the reasons why philosophers have paid little attention to the problem of self-other discrimination, and investigate the conditions under which this problem can arise. In section 2, we will concentrate on the identification of the self as a physical body. In section 3, we will present evidence that we are able to perceive the intentions of others. This will involve the idea of shared representations of actions, the nature and format of which we will

Address for correspondence: Institut des Sciences Cognitives, 67 Boulevard Pinel, 69675, Bron, France
Email: jeannerod@isc.cnrs.fr

Address for correspondence: Institut Jean Nicod, CNRS-EHESS-ENS, 1 bis avenue de Lowendal, 75007 Paris, France
Email: pacherie@ehess.fr

discuss. We consider possible mechanisms underlying self-other discrimination and the attribution of actions and intentions to their authors. Two hypotheses will be presented for self-identification: the central monitoring hypothesis and the simulation hypothesis. In section 4, we will investigate how impairments of these mechanisms may underlie some pathological conditions involving misattributions of actions and intentions. The final section summarizes our findings and our arguments for the claim that the problem of self-identification does arise, not just for the self as object, but also for the self as subject.

1. Self-other distinction: Is there really a problem?

A number of contemporary thinkers acknowledge that when I judge 'I think X', I may be mistaken about X and thus that the mind is not wholly transparent of itself. But they maintain, with Descartes, that when I judge 'I think X', I cannot be mistaken about who the subject of the thought is. Judgements of this kind are said to be immune to error through misidentification, because they are identification-free. As Shoemaker puts it: 'One refers to oneself, and ascribes properties to oneself without having to identify as oneself anything that is presented to one' (1996, p. 196). The problem of self-identification can therefore only arise for those forms of self-ascription that are not identification-free. We claim that the self-ascription of agency is not identification-free.

For the Cartesian, the problems of the external world and of other minds were the price of certainty about the self. Kant, on the other hand, argued that a condition of self-consciousness is a capacity to distinguish between self and world, to think of at least some of one's experiences as experiences of objects existing independently of being experienced. Strawson then (1959, 1966) argued that, for one's experience to meet this condition, one must be aware of oneself as a physical object located in space, whose experiences are determined at least in part by where in space one is. A further thought, implicit in Strawson's argument, is that awareness of oneself as a physical entity located in space and with experiences determined at least in part by where in space they are located, requires the capacity to actively move through space. If our spatial viewpoint were fixed, we would not have the idea of places other than those we are perceptually related to; we would not have the idea of things existing unperceived in parts of space where we are not located. And if we were only passively moved though space, we wouldn't be in a position to distinguish between changes in our experiences that are simply a consequence of our having changed places—position-changes—and changes in our experience corresponding to actual changes in the world—state changes. As argued by Poincaré (1907), what distinguishes the two kinds of experiences is that experiences of position-changes, but not experiences of state-changes, can normally be reversed by making compensatory movements that replace us in the initial situation. According to him, it is by trial and error that we come to

learn which voluntary movements can compensate for which changes and thus acquire a representation of space. If we were only passively moved, there would be no guarantee that we would have a sufficient experience of reversibility to construct a representation of space. Indeed, an individual who spent all his life being passively moved in a straight line would not even experience reversibility.

In sum, self-consciousness requires the objectivity condition, which requires the agency condition, i.e. awareness of oneself as an agent.[1] Once it is accepted that self-consciousness involves being able to think or at least be sensibly aware of oneself as a physical object among others, one form of the problem of self-identification can arise. The bodily self is one object of experience among many. It therefore makes sense to ask: how do I identify my bodily self? How do I discriminate it from those sometimes very similar human bodies I also perceive[2]?

Although it may assuage skeptical doubts about the existence of the external world, the Kantian move leaves intact the problem of other minds. A further inferential move is needed, it seems, to get from the existence of other bodies to the existence of other minds. Cartesianism assumes a fundamental asymmetry between the way we know of our own mental states and the way we know the mental states of others. As long as this asymmetry is in place, the self-other distinction cannot arise as a serious philosophical problem. Yet this makes problematic the semantics of mental state ascriptions: how can the meanings of our mental terms and concepts be univocal if the criteria we apply in the two cases are so different? Wittgenstein (1953) and a number of philosophers after him argued that the criteria for the application of mental terms must be public, hence these terms get their meaning from the publicly observable behavior of agents. This shift towards a third-person conception of mentality has counterintuitive consequences however, for it seems to imply that we know our own mental states on the basis of our behavior.

We face a dilemma: we stick to the traditional asymmetry with its implausible semantic consequences, or we adopt a third person perspective, at the cost of an

---

[1]   For an in-depth discussion of the various versions of the objectivity argument, see Cassam (1997). One important distinction stressed by Cassam is between a concept version of the objectivity argument that requires that one think of oneself as a physical object among others and an intuitive version of this argument that requires that one be intuitively or sensibly aware of oneself as a physical subject. There are similarly various versions of the agency argument, i.e. the argument that links the experience of an objective world with the experience of agency. The argument briefly outlined here and developed in Pacherie (1997) is one version of it. For other versions, see for instance Baldwin (1995) and Russell (1995).

[2]   In *The Blue Book*, Wittgenstein (1958) distinguishes between the self as object and the self as subject and gives as an example of a mistaken identification of the self as object the case of a man in an accident who feels a pain in his arm, sees his neighbour's broken arm at his side and mistakenly takes it to be his arm. We fully agree with Wittgenstein that error through misidentification is possible for the self as object, but we argue furthermore, against his view, that error through misidentification can also arise for the self as subject. Two forms of the problem of self-identification must therefore be distinguished.

implausible conception of our knowledge of our own minds. One way out involves rejecting the radical dualism of the Cartesian picture: mental states are internal states whose intrinsic mental character is independent of anything external, while human bodies are merely material objects and their behavior is intrinsically non-intentional. On this view, what confers intentional properties on behavior are its inner mental accompaniments and causes. Nothing intrinsic distinguishes a mere bodily happening from a piece of intentional behavior; the difference is one of causal antecedents. Since internal mental causes cannot be directly observed, they must be inferred and it is always possible that the inference be incorrect.

One way to reject this picture, championed by McDowell[3], involves taking the concept of a 'human being' as basic, 'as a seamless whole of whose unity we ought not to have allowed ourselves to lose sight in the first place' (McDowell, 1982, p. 470). Behavior and mentality are then much more integrated and, in favorable circumstances, the mental states of others can themselves be objects of experience. We offer a limited defense of this approach, concentrating on one type of mental states, namely intentions.[4] The claim we want to defend is that intentions can be, at least in some circumstances, transparent in behavior, that they can be directly observed and do not have to be inferred from behavioral proxies. But our defense will be rather different from McDowell's. We will focus on the links between perception and action, and present some empirical evidence that intentions can be perceived.

First, we need to be clear about what, on our account, it is to perceive an intention. In the theory of perception, it is standard to distinguish primary, secondary and reflexive awareness. When I see a tree, I am primarily aware of the tree, only secondarily aware of myself as in front of the tree, and only reflexively aware of myself as in a certain perceptual state, namely seeing the tree. We claim that it is like this with the perception of intention: when Mary watches John open the door, she is primarily aware of an intention to open the door, rather than being primarily aware that John intends to open the door. Similarly, when Mary herself intends to open the door, she is primarily aware of an intention to open the door, rather than being primarily aware that she herself intends to open the door. Let us call this awareness of an unattributed or 'naked'[5] intention. Then our argument will proceed as follows:

(1)   The intentions of others can be directly perceived;
(2)   Our primary awareness of intentions, either ours or those of others, is as naked intentions;

---

[3]   Note that, according to McDowell, this was also Wittgenstein's view. The idea that the notion of a 'human being' or a 'person' should be taken as primitive is also present in Strawson (1959).

[4]   Another category of mental states for which McDowell's claim appears plausible are emotions. Whether this claim can be sustained for all types of mental states, including, for instance, beliefs, appears more doubtful. In any case, we will not tackle these issues here but will restrict ourselves to intentions.

[5]   By analogy with the phrase 'naked infinitive'.

(3)   The channels that provide the collateral information needed to deter-
mine who the author of the intention is are not entirely reliable and
cannot be a source of identification–free first–person knowledge.

In sections 2, 3 and 4 we will review empirical results relevant to the assessment
of conditions (1)–(3). In section 5, we will argue that these conditions can and do
indeed obtain.

2. Bodily self–identification

According to standard philosophical wisdom, error through misidentification is
possible for the self as object, but not for the self as subject. We will argue that both
types of misidentifications are actually possible and that identification of the self as
subject and identification of the self as object are often interdependent. It is
therefore important to consider how we identify or recognize one physical body
as our own body (identification of the self as object). First, the matching of visual,
tactile and proprioceptive signals originating from the same body parts contributes
to an intermodal sensory image of the body. Second, our body is also an acting
body and the matching of one's intentions with the bodily effects of self-generated
actions contributes as well to our sense of bodily ownership.
   It is important to distinguish between identifying a given body (or body part) as
one's own (bodily self-identification or recognition) and recognizing our body or
one of our body parts as having a certain property (bodily self-predication).
Deciding whether a given arm is or isn't mine is a matter of self-identification
but deciding whether my leg is straight or bent is a matter of bodily self-
predication. Proprioception and passive touch are *de facto* immune to error through
misidentification. Although it seems conceivable that one could be hooked up to
other bodies in such a way that one had proprioceptive and tactile access to their
states as well, in the actual world proprioception and passive touch only carry
information about one's own body.[6] However, errors of bodily self-predication
can arise, as with the question whether my leg is really bent or whether tactile
stimulation is really applied to my hand.
   In contrast, vision can carry information both about one's own body and the
body of others. It makes sense therefore not just to ask whether the leg I see bent is
really bent but also whether this is my leg or not. As a consequence, visual
perception can mislead us in two different ways: one may mistakenly think that

---

[6]   The case of passive touch is slightly more complicated than that of proprioception, for when
we feel a tactile simulation we are aware not just of our body as being touched but of
something else that touches the body. Our point though, is that although we might be
mistaken as to the identity of the object that touches us, it is not possible to be mistaken as
to who it is that is being touched. In that sense, passive touch is also an identification-free form
of awareness of one's body.

one's leg is bent as a result of having seen a bent leg, and one may mistakenly think that a seen leg is one's own leg. So proprioception is susceptible to error only of self-predication while vision is subject to this error and also to error of self-identification. Yet, despite this twofold fallibility, many experimental results stress a prevalent role for vision over other senses in bodily self-predication and self-recognition. For example, experiments show that where we feel (proprioceptively) our hand to be is determined by where we see it. Optical distortion of the visually perceived position of a limb with respect to its felt position (e.g. by wearing laterally displacing prisms) produces no alteration of the sense of ownership: we do not come to doubt whether the hand we see is really our hand, instead the position sense is recalibrated to conform with the visual information (Harris, 1965). This primacy of vision was confirmed in experiments using a rubber arm. Botvinick and Cohen (1998) positioned a realistic rubber arm in front of subjects, while their real arm, hidden by a screen, was placed aside: tactile stimulation was applied simultaneously to the real and the rubber arms. After some time, the subjects experienced an illusion in which they felt the touch at the locus of the rubber arm (that they could see), not of their real (hidden) arm. In other words, the tactile stimulus was felt where it was seen, at the expense of a distortion of the felt position of the real arm. In addition, subjects spontaneously reported experiencing a clear sense of ownership for the rubber arm. According to other authors who replicated this experiment, the impression of displacement of the tactile stimulus and the illusion of ownership disappear if the rubber arm is not properly aligned with the subject's body (Farné *et al.*, 2000).

Let us examine in more detail what is going on in this experiment. Note, first, that three types of sensory information, visual, tactile and proprioceptive, are in play. It is important to differentiate the respective contributions of touch and proprioception, which the verb 'to feel' conflates. Tactile information, by itself, only tells us which body part is stimulated, here the arm, but it doesn't tell us about the position of the arm relative to the trunk. This latter information is given by proprioception. In general, vision, however, gives us information both about the position of our body parts and about which body part is stimulated. The sense of ownership for the rubber arm happens as a result of the congruence of visual and tactile information, both telling the subject that his arm is stimulated, together with visual information about the alignment of the arm with respect to the subject's body, and despite the incongruence of the visual and the proprioceptive information about the arm's position. The relationship between bodily property-predication and bodily self-attribution is therefore not completely straightforward. On the one hand, the congruence of tactually based and visually based property-predication (stimulation is applied to the arm) appears to contribute to the sense of ownership for the arm. On the other hand, the incongruence between proprioceptive and visual information about the arm position does not appear to be taken as evidence against ownership of the arm.

However, the exploitation and comparison of sensory cues is not our only basis for identifying a seen body or seen body parts as our own. The self is not just a

sentient self; it is most of the time an acting self. Action-related cues also make an important contribution to the sense of body-ownership. Moreover, the working of those action-related cues also shows the close interplay between a sense of owner-ship and a sense of agency. Body parts are moving with respect to one another and with respect to external objects as the result of intentional actions. It is common experience that our actions are readily self-attributed as a consequence of a normally perfect correlation between their expected effects and the flow of resulting (visual and proprioceptive) stimulation. This matching process provides the agent of an action with the sense both that the acting body he sees is his own (sense of ownership) and that he is causing the action (sense of agency). As is illustrated by the famous mirror scene in the Marx Brothers' film, *Duck Soup*[7], an efficient means for determining whether the body or the body parts we see are ours is to move: if the image of a body I see in a mirror in front of me moves when I move and the two movements are congruent, then the image must be an image of my body.

Again, however, situations may arise where this attribution becomes less than obvious. We will now describe experiments where congruence was measured and manipulated. Their results point to the main cues used by normal subjects to determine both ownership and authorship.

2.1 The Nielsen's paradigm

Nielsen (1963) designed an experiment where an alien hand is substituted for the subject's own; the subject is then presented with hand movements departing from his own movements. At which point will the subject become aware of the substitution? In the experiment, subjects were requested to draw a straight line on a piece of paper placed in front of them. In one condition, they saw their own hand doing the task; in another condition, they were shown the experimenter's hand doing the same thing at about the same time. When unbeknownst to them, the experimenter's hand drew a line that deviated from the requested direction, the subjects tended to shift the trajectory of their own (unseen) hand in the opposite direction, so as to fulfil the instruction they had received to draw a straight line. The subjects were unaware of this deviation: when shown their deviant performance, they tried to explain it by such factors as fatigue or inattention.

Fourneret and Jeannerod (1998) devised a slightly different version of Nielsen's experiment. Subjects were instructed to move a stylus with their unseen hand to a visual target: only the trajectory of the stylus was visible as a line on a computer screen, superimposed on the hand movement. In some trials, a directional bias (2°, 5° or 10° to the right or to the left) was introduced electronically, such that the visible trajectory no longer corresponded to that of the hand. In order to reach the target, the hand held stylus had to be moved in a direction opposite to the bias. In other words, although the line on the computer screen appeared to be directed to

---

[7]   In this scene a mirror breaks, and Harpo imitates Groucho, move for move, to postpone the discovery of the broken mirror.

the target location, the hand movement was directed in a different direction. At the end of each trial, subjects were asked in which direction they thought their hand had moved by indicating the line corresponding to their estimated direction on a chart presenting lines oriented in different directions.

This experiment revealed several important points. First, subjects accurately corrected for the bias in tracing a line that appeared visually to be directed to the target. This resulted from an automatic adjustment of their hand movements in a direction opposite to the bias. Second, subjects tended to ignore the veridical trajectory of their hand in making a conscious judgement about the direction of their hand: instead, they tended to adhere to the direction seen on the screen and based their report on visual cues, thus ignoring non-visual (e.g. proprioceptive) cues. The general idea suggested by this result is that the visuomotor system is able to appropriately use information for producing accurate corrections to reach a target, but that this information is not accessed consciously.

The problem now would be to determine the threshold for conscious access to the correction generation mechanism. It is conceivable that the automatic system should not be able to produce corrections for larger and larger biases and should break down at some point, with the consequence for the subjects of noticing errors in their attempts at reaching the target. The effects of pushing the automatic system to its limits was thus examined in a further experiment using the same apparatus (Slachewsky *et al.*, 2001). The bias that, in the previous experiment, was randomly presented and was limited to a maximum angle of 10° to the right or to the left, was progressively increased from trial to trial. When the bias amounted to a mean value of about 14°, subjects changed strategy and began to use conscious monitoring of their hand movement to correct for the bias and to reach the target. In other words, the discrepancy between the seen trajectory and the felt trajectory became too large to be automatically corrected, and the failure of these corrections was compensated by conscious deviations of the hand movement in the appropriate direction.

Let us take stock. The purpose of these experiments was to investigate the types of cues and congruence among them used by normal subjects to determine both bodily ownership and the sense of agency. These cues are of two main kinds: on the one hand, the awareness a subject has of what he intends to do—what we may call the awareness of willing—and, on the other, sensory cues—in the present case visual and proprioceptive information about the trajectory of one's arm. One should also distinguish three aspects or components of the experience of willing: one may be aware that one intends to move (that-experience), one may be aware of the goal of the movement, here reaching for a certain target (what-experience), and finally one may be aware of more specific parameters of the movements one intends, e.g. moving one's arm at a certain speed along a certain trajectory (how-experience).

What the experiments described in the previous paragraphs show is, first, that the how-component of the experience of willing is very coarse-grained: we are aware of some global parameters of the intended movements, but we do not experience willing specific muscle contractions, joint torques, or movement

velocities that are also necessary to achieve the goal. Our conscious access to our motor processes is thus extremely limited. Second, the comparison processes that take place during execution are mostly unconscious. What is available for comparison with the experience of willing is mostly visual information pertaining to the achievement or non-achievement of the goal. It is only when the discrepancy between visual and proprioceptive signals becomes severe that we become aware of the proprioceptive signals. Thus, typically, when there is no serious discrepancy, the experience of willing gives rise both to the experience of acting—I am causing this movement—and to the experience of bodily ownership—the effector that is producing the movement is mine—when the agent has matching visual information. This suggests that both the experience of willing and the experience of acting are predominantly what-experiences and that we are mostly unaware of the how-aspects of an action. It may be however, that proprioceptive signals contribute to our experience of acting not by making us aware of how exactly movements are performed, but simply by contributing to our awareness that we are moving (see Jeannerod, 2003). In other words, it may be that they function mainly as the reafferent sensory counterparts[8] of the internally generated sense of effort and that their main contribution is to our experience *that* we are acting rather than to an experience of *how* we are acting.

Clearly then, the experience of acting is not infallible. However, it is interesting to distinguish several ways in which it can go astray, and hence several kinds of illusions it may foster. First, there may be cases where a subject is deprived of sensory reafferences, as in the famous case reported by William James (1950) where a patient with an anaesthetized arm is blindfolded and then asked to raise his arm, while, unbeknownst to him, his arm is prevented to move. In the absence of sensory reafferences, and if he has an explanation for this absence, a subject may well interpret an experience of willing as an experience of acting. In this case, the subject is under the illusion that he is acting: he feels as though he is doing something when he is doing nothing. In the experiment by Fourneret and Jeannerod (1998), the subjects experience another kind of illusion. Their experience of acting is non-veridical not in that it tells them that they are acting when they are not, nor in that it tells them that they are achieving their goal (moving towards the visual target) when they are not, but in that it tells them that the way they do it is by moving their hand in a certain direction. In other words, they feel as though they are doing something in a certain way when actually they are doing it in a different way. Finally, the experience of acting can also give rise to an illusory sense of bodily ownership, when as in Nielsen's experiment, the matching of the willed movements with the seen movements of the hand is good enough to be taken as evidence by the subject that the seen hand is his own. Indeed, simply looking at a moving limb superimposed on one's own limb creates a strong impression of having willed this movement and of being its author.

---

[8]   That is, the sensory signals generated by the movements produced.

Observations have been reported in amputees who experience having a phantom limb. When their valid limb is visually transposed (by way of mirrors) to the amputated side, and when, on being aked to try to move both arms simultaneously, they produce movements with their valid limb, they experience a strong feeling of voluntary movement of the phantom limb. The same happens if the visually transposed limb is that of an experimenter (Ramachandran and Rogers-Ramachandran, 1996). These observations are particularly interesting in that the illusion experienced by the amputees is twofold: their experience of acting tells them that they are moving their (amputated) arm when obviously they are not and that the arm they see moving is theirs. As argued by Wegner (2002), these observations reflect a tendency to perceive oneself as causally effective. When what we see corresponds to what we were trying to do, or more precisely to what we were aware we were trying to do, we tend to attribute the cause to ourselves. This tendency might account for the frequent observation of over-attribution to the self in subsequent experiments (see below).

2.2 Self-identification in the social context

Although this series of experiments based on Nielsen's paradigm dealt with the recognition of action and with the degree of awareness of his own movements a subject can have, it said little about the cues that a subject can use for attributing this action to its origin. One needs to distinguish between fallibility with respect to what action one is performing and fallibility with respect to whether it is you or not who are performing an action, i.e. fallibility with respect to action attribution. Obviously, the two forms of fallibility are not unrelated; whether an action we observe is our own presumably depends at least in part on whether we think the movement we observe corresponds to the movement we intended to perform. However, action attribution can be seen as a further step in self-identification, where what is at stake is identification of one's own self among other selves. This problem was absent from the above experiments by Fourneret and his colleagues, where the subject was confronted with his own performance without any hint that the movements he saw could have been produced by another self.

   In order to explore the factors involved in attributing an action to the self or to the other, a situation was created where the subjects were shown the movements of hands of an uncertain origin, that is, hands that could equally likely belong to themselves or to someone else (van den Bos and Jeannerod, 2002). The experimental situation involved the simultaneous presentation of two hands, one of which was the subject's hand, the other being an alien hand. Subjects were instructed to explicitly determine whether or not they were the authors of the hand movements they saw. The images of the two hands filmed by a TV camera were displayed on a computer screen. A program made it possible to rotate the image displayed on the screen, so that the subject could see the hands where they were in reality (0° rotation), or rotated by 180° or 90° with respect to their actual position. In addition, subjects were instructed on some trials to move their fingers

during presentation of the hands. On the trials where the subject was instructed to make a movement, the experimenter would either make the same or the alternative movement. At the end of the trial, the hands disappeared and an arrow appeared on the screen pointing to a position where one of the two hands had been. Subjects had to say 'yes' if they thought that the hand indicated by the arrow was their hand and 'no' if they thought it wasn't.

This experiment first allowed us to study the role of the apparent positions of the hands in self-recognition. When the two hands visually appeared at the loci corresponding to their real positions, subjects showed relatively little difficulty in recognizing their own hand. However, when the apparent locations of the hands were interchanged with respect to reality, they made attribution errors. This confirms that the contingency between visual and proprioceptive signals plays a role in self-recognition.

The most critical factor for correct attribution was the presence of finger movements. When finger movements were present and these movements were clearly attributable to the self (i.e., they differed from those of the experimenter), no attribution errors were found. Accurate self-recognition was possible for all orientations of the display, including the 180° rotation. In other words, when distinctive movements are available, subjects tend to recognize actions, not just hands. By contrast, when the two hands performed the same movements or no movements, the contingency between visual and proprioceptive signals could not be used and the error rate increased as a function of the degree of rotation. Finally, when they made errors, subjects tended to misattribute the indicated hand more often to themselves than to the other. This finding of over-attribution to the self, which was observed in several other experiments (e.g. Daprati *et al.*, 1997; Franck *et al.*, 2001), will have important implications for understanding the pattern of misattribution in pathological conditions (see below).

Let us conclude this section with two general remarks. First, the main finding common to the experiments reported in this section is that although several sources of information are involved in bodily self-predication, bodily self-identification and action attribution, their respective contributions do not carry the same weight. Visual information very often dominates over both information delivered by other sensory modalities and information carried by central signals. In bodily self-predication, we tend to ascribe to our body (and body parts) the properties it is seen to have, even when our visual perception is non–veridical and conflicts with information from other modalities. This in turn may lead to erroneous bodily self-identification as when one feels a sense of ownership for a rubber arm. The dominance of visual information is not confined to the sensory sphere; it also extends to the sphere of action. If he sees his hand moving in the intended direction, a subject will tend to think that this is what he is doing even though proprioceptive and central motor signals say otherwise. Or to be more precise, in the experience of willing, information about the goal plays a much greater role than information about motor means and, in the perceptual identification of one's body and its properties, visual information dominates over information from other

sensory modalities. When the two coincide, when what you see is what you intended to do, what they say is not easily undermined. Only when serious discrepancies with information from other modalities arise, can these 'recessive' modalities make themselves heard, so to speak.

Second, as we pointed out at the beginning of this section, proprioception and passive touch only carry information about one's own body, whereas vision can carry information both about one's own body and the body of others. Thus, bodily self-identification (and misidentification) is typically visual.[9] Three kinds of visual characteristics of a body or body part can play a role in its identification as one's own or not. Its general appearance (size, shape, texture, color, and so forth), its movements and its position in the visual field. In the experiments we reported, visual appearance was neutralized by having the subjects wear gloves. The remaining cues were movements and position. The experiment in van den Bos and Jeannerod (2002) shows the importance of movements as a negative cue. However, when this cue is in turn neutralized (no movements are performed or the same movements are performed by both subject and experimenter), position becomes crucial. How is visual information about position used in bodily identification? Here it may help to distinguish between the body image, conceived as a long-term representation of the way our various bodily segments are connected and of the degrees of freedom of our joints, and the body schema, the short-term, constantly updated, representation of the position of our limbs at a given time.[10] We may say that the body image translates visually as a representation of the range of positions the various body parts can occupy in egocentric visual space and of the ways they should be spatially related one to another. For instance, to see our feet we typically have to look down, to see our right shoulder we should look both to the right and somewhat down, our right arm should be seen in a position compatible with its being attached to the right shoulder, our back and most of our face cannot be seen without using mirrors. When we see a body or part of a body and the only relevant information is information about position, how do we decide whether it is ours? The results from van den Bos and Jeannerod (2002) suggest that in such a case compatibility with the body image rather than the more specific body schema is the crucial factor.

3. The nature of the mechanisms of self–identification

There are several ways of conceiving the mechanisms involved in self-identification. As we shall see in this section, however, most if not all hypotheses about these mechanisms refer to one aspect of self-identification, namely, action recognition

---

[9] Note though that audition and haptic perception (active touch) may also involve self-identification.

[10] For more detailed characterizations of the distinction between body image and body schema, see Paillard (1982, 1999); Cole and Paillard (1995), and Gallagher (1986).

understood as identification of oneself as the author of the action. This predominance of action recognition over other possible cues for self-identification, which is clearly highlighted by the above experiments, opens a number of empirical and theoretical questions. In this section, two empirically based hypotheses of action recognition will be described. One hypothesis relies on the idea that executed actions generate signals that are centrally monitored and compared: action recognition arises as the outcome of this comparison (the central monitoring theory). The other hypothesis relies on the idea that actions, whether or not they come to execution, are centrally simulated by the neural network, and that this simulation is the basis for action recognition and attribution (the simulation theory). These two theories will appear to be largely complementary, with the simulation hypothesis supplementing the central monitoring hypothesis, and, we will argue, overcoming some of its limitations.

3.1 The central monitoring theory of action recognition

The first theory to be considered is that of a monitoring of central and peripheral signals arising as a consequence of the execution of an action. This theory holds that the (central) efferent signals at the origin of an action are matched with those which result from its execution (the reafferent signals) and that this comparison provides cues about where and when the action originated. Let us first rephrase its basic principles. The original idea, inherited from the cybernetic era, and deeply rooted in physiological thinking, is that each time the motor centers generate an outflow signal for producing a movement, a 'copy' of this command (the 'efference copy') is retained. The reafference inflow signals (e.g. visual, proprioceptive) generated by the movement are compared with the copy so that, if a mismatch is recorded between the two types of signals, the state of the motor centers can be updated (Sperry, 1950; von Holst and Mittelstaedt, 1950).

   More recent descriptions of this mechanism assume the existence of 'internal models' where the desired state of the system is represented. Internal models can predict the sensory consequences that should result from the execution of the motor command. In a self-generated movement, predicted consequences are compared with the actual sensory feedback (Wolpert *et al.*, 1995). This model of the control of action can be directly applied to the problem of self-identification. Self-identification could be based on the degree of concordance between the predicted effects of an action and its actual sensory consequences. If the concordance is complete, then the executed action corresponds to the desired one and can be automatically attributed to the self.

   Neuroimaging techniques, which allow monitoring brain activity in normal subjects, have been used in situations where the degree of concordance between the various signals produced by a self-generated movement can be manipulated. Consider for example a situation where a subject executes a movement and receives a discordant visual feedback from this movement: the subject will experience a conflict between his intention and his senses, i.e. the anticipation of seeing a certain effect as the result of his intention (e.g. his hand moving in a certain

direction) will be contradicted by the actual feedback he will observe (several such situations have already been described section 2). A new experiment using this paradigm and where, in addition, brain activity was monitored by PET, was undertaken by Farrer *et al.* (2003). In this study, the degree of concordance of the visual feedback provided to the subjects about their own movements was varied. Visual feedback could be either fully congruent with the executed movements (i.e. the subject saw exactly what he did), or distorted up to the point where the seen movements were completely unrelated to the executed ones. Thus, in the congruent condition, the subjects were likely to feel in full control of their own movements, whereas in the maximally distorted condition, they were likely to feel that they were not in control, but rather being overridden by the movements of another agent. To achieve this, subjects were instructed to continuously move a joystick with their right hand. The hand and the joystick were hidden from subjects' view. Instead, the subjects saw the electronically reconstructed image of a hand holding a joystick appearing at the precise location of their own hand. When the subject moved, the electronic hand also moved by the same amount and in the same direction: subjects rapidly became acquainted to this situation and felt the movements of the electronic hand as their own. Distortions were introduced in this system, such that the direction of the movements seen by the subjects were rotated with respect to the direction of those they actually performed. A graded discordance was produced by using a $25°$ rotation, a $50°$ rotation and finally a situation where the movements appearing on the screen had no relation with those of the subjects (they were actually produced by an experimenter). Subjects were instructed to concentrate on their own feelings of whether they felt in control of the movements they saw.

The results of this experiment demonstrate that a decreasing feeling of control was associated with changes in neural activity in several specific brain areas. For the purpose of this paper, we will concentrate on one of these areas, the inferior parietal lobule. The reason for this choice is twofold: first, the inferior parietal lobule is an associative region where signals relevant to our present discussion (e.g. visuomotor and proprioceptive signals, signals related to motor commands), are likely to be processed; and second, lesions of the inferior parietal lobule produce pathological effects which clearly indicate its involvement in action recognition (see Bisiach and Berti, 1987 for a review of these effects). Thus, the degree of activation in the inferior parietal lobule (more particularly on the right side) increased when the discordance between central signals arising from the motor intention and visual and kinesthetic signals arising from movement execution became larger. A possible explanation for this result is that the mismatch between normally congruent sets of movement-related signals increased the level of processing of these signals and, consequently, produced a larger metabolic activity.

The results from the above experiment seem to give some support to the central monitoring theory insofar as they show that neural activity in a specific brain area correlates with the feeling of being in control of the action—an increased activation being associated with a decreasing feeling of control. Yet, it remains to

be seen whether the central monitoring theory can provide a general solution to the problem of self-identification. This theory appeals to neural signals produced by the subject's motor activity and postulates that these signals are used for comparing the end result with the internal model of the action. As we shall now see, however, there are many situations where an action representation is formed but no movement is actually executed. In such situations, there are no output signals to the muscles, no reafferent (e.g. visual) signals from the outside world, no proprioceptive signals and, therefore, no possibility for comparing execution with a desired output.

3.2 Imagining and observing actions: simulation and the limits of the central monitoring hypothesis

The idea that an action should be executed for self-identification to operate, as postulated by the central monitoring theory, is indeed problematic. The existence of overt behavior is not a prerequisite for self-attribution. When we imagine ourselves performing an action (the paradigmatic case of motor imagery), we unambiguously attribute this action to ourselves. On the other hand, when we observe an action performed by another person, we unambiguously attribute it to that person. In this section, we will describe recent research on motor imagery and action observation in order to validate the notion of *covert action*. We will hypothesize that covert actions are simulated actions that do not involve contraction of muscles and therefore, that remain invisible from the outside. We will argue that, whereas every overtly executed action implies the existence of a covert stage, a covert action does not necessarily turn into an overt action (Jeannerod, 1994).

Interest in the field of mental imagery arose more than a century ago in the context of hypnosis. To explain hypnotic phenomena, authors of the time (e.g. Binet, 1886) claimed that mental images in general resulted from excitation of the same cerebral centers as the corresponding actual sensation. In the domain of motor images, it was remarked that the state of the motor centers influenced the possibility to generate a motor image: for example, it was shown to be impossible for a subject to generate the image of pronouncing the letter /b/ if he kept his mouth wide open: this is because, it was conjectured, the motor system cannot be engaged in two contradictory actions at the same time. More recently, imagined actions have become a major tool for the study of representational aspects of action. Behavioral findings revealed that motor images retain the same temporal characteristics as the corresponding real action when it comes to execution. For example, it takes the same time to walk mentally to a prespecified target as it takes to actually walk to the same place (Decety *et al.*, 1989). Similarly, temporal regularities that are observed in executed actions, such as the classical speed–accuracy tradeoff, are retained in their covert counterparts (Sirigu *et al.*, 1996).

Explicit motor imagery, however, is only one of the forms an action representation can take. Conscious access is not a distinctive characteristic of covert actions: on the contrary, the representation of the action may remain entirely unknown to the agent and its content only become visible through the motor decision the agent

takes. Consider for example the situation where a subject is simply requested to make an estimate about the feasibility of an action, e.g. to determine the feasibility of grasping an object placed at different orientations: the time it takes to give the response is a function of the object's orientation, suggesting that the arm has to be mentally moved to an appropriate position before the response can be given. Indeed, the time to make this estimate is closely similar to the time it takes to actually reach and grasp an object placed at the same orientation (Frak *et al.*, 2001; see also Parsons, 1994).

Another covert form of action is action observation. To observe an action with the purpose of understanding it or with the purpose of understanding the intention behind it is by no means a passive process. This idea was already present a century ago under the concept of empathy (e.g. Lipps, 1903). Empathy expresses the possibility that we understand other people's behavior (e.g. their actions, their facial expressions, etc) because we attempt to replicate and simulate their mental activity. More recently, the idea that mental simulation plays an essential role in our understanding of the mind has been vigorously argued for by a number of philosophers and psychologists. Indeed, in the last two decades, a complex interdisciplinary debate on the nature of folk-psychology has been going on between theory theorists (Leslie, 1994; Gopnik and Meltzoff, 1997; Stich and Nicholls, 1995) and simulation theorists (Goldman, 1995; Heal, 1994, 1995, 1996; Gordon, 1995a, 1995b, 1996; Harris, 1995). Theory theorists claim that our capacity to attribute mental states and predict behavior is best accounted for by possession and use of a theory of mind—a body of knowledge about cognition and motivation equally applicable to oneself and to others. Simulation theorists maintain instead that it is better thought of as an ability to simulate—to use our own mental processes to simulate those of other agents. It should also be noted that there are many possible versions of both theory theory and simulation theory.

There are weaker and stronger versions of simulation theory varying along the following five main dimensions of variability, the first three of which are described in Goldman (2002). The first is frequency of use: one may say that simulation is used in all cases, take it to be the default procedure that may at times be overridden by theoretical considerations, or think of it as employed only occasionally. A second dimension of variation is what Goldman calls the source of simulation. Is simulation the primary, root form of interpersonal mentalization or is it merely a useful short-cut to theoretical reasoning in the fashion theory theorists conceive of it? A third dimension of variability is the scope of simulation. Is simulation used for the attribution of all kinds of mental states or only for a restricted range of them? In other words, can all the mechanisms involved in the generation of mental states be deployed in simulatory fashion or only a subset of them? A fourth dimension of variability is explicitness. Simulation can be conceived as the explicit, conscious imaginary enactment of the mental states and processes of others. It can be considered as a subpersonal process unfolding automatically and without conscious control. Or it can be thought of as a hybrid of implicit and explicit simulation. Finally, some simulation theorists want to limit simulation theory to being a theory

of propositional states (Heal, 1996) whereas others, like Goldman, extend simulation to include mental imagery.

Here, our focus is on simulation in the domain of action and although we think that simulation is not restricted to just this domain, we will not enter a discussion of what its exact scope is. As far as the understanding of action is concerned, we regard simulation as the default procedure that can occasionally be supplemented or overridden by theoretical considerations. We also believe that simulation is the root form of interpersonal mentalization and that it is best conceived as a hybrid of explicit and implicit processes, with subpersonal neural simulation serving as a basis for explicit mental simulation.

From an empirical standpoint, adherence to simulation theory in the domain of action commits us to three related ideas. Simulation would not be a very useful tool if we were not able to identify plausible targets for simulation, to somehow distinguish between intentional and non-intentional behavior. The plausibility of a simulationist stance towards actions therefore rests on the idea that that intentional bodily behavior has distinctive intrinsic characteristics. Furthermore, it is not enough that it have such characteristics, we must also be perceptually sensitive to them. The existence of distinctive characteristics of intentional behavior and their detectability are thus two preconditions for simulation. But that these preconditions obtain does not guarantee that we are indeed simulating observed actions. For talk of simulation to be legitimated, the processes and mechanisms used for the preparation of our own actions should also be involved when observing actions performed by other agents; in other words, an observed action should activate, in the observer's brain, the same mechanisms that would be activated, were that action intended by the observer (Gallese and Goldman, 1998).

We now have ample empirical evidence that these three conditions do indeed obtain. Intentional behavior has distinctive observable properties, a distinctive kinematics and a dynamics that bears systematic relations to features of the situation in a way that non-intentional behavior does not. Perception of biological motion has been extensively studied by psychologists using the point-light technique developed by Johansson (1973), where an actor is filmed in a dark room with lights attached to his main joints and where as a consequence only kinematic information is available to the perceiver. In his pioneering studies, Johansson filmed movement patterns of walking, cycling, climbing, and dancing by the point-light stimulus agents. He showed that subjects who could not identify these actions when presented with still frames where lights were stationary and only information about joint positions was available would quickly and reliably recognize them as soon as the lights were moving. Using the same experimental paradigm, other investigators further demonstrated that subjects were able to identify friends by their gait (Cutting and Koslowsky, 1977) as well as the gender of walking person (Koslowsky and Cutting, 1977) and that people can recognize themselves more easily than others (Beardsworth and Buckner, 1981). Furthermore, perceptual sensitivity to human motion seems to be innate or to develop very early (Bertenthal *et al.*, 1984). These results suggest that visual perception of

actions is influenced by implicit knowledge of the motor rules involved in their production. Other studies have also shown that this knowledge can be exploited by the visual system not just for the recognition of action, but also for the recognition of intentions. Thus, in a series of experiments using the point-light technique, Runeson and Frykholm (1983) have investigated the perception of intentions in object-oriented actions. For instance, when an agent has to lift a box, her expectations regarding the weight of the box lead her to make specific anticipatory postural adjustments; when the weight is not as expected postural readjustments take place. These investigators have shown that subjects observing point-like displays of such actions are able to determine what the actor expects the weight to be and whether her expectations are correct. Furthermore, they are able to detect attempts at deception. If an actor pretends, say, that a suitcase he is carrying is heavier than it actually is, his movements will have a non-natural kinematics that can be detected by observers. Another series of experiments (Orliaguet *et al.*, 1997; Kandel *et al.*, 2000) investigated motor anticipation in the perception of handwriting gestures. For instance, when writing an l, the kinematics of the graphic production is influenced by anticipation of the next letter to be written and will differ according to whether the l is to be followed by another l, an i or a n. It was shown that the visual system can exploit these differences to predict the second letter.

Finally, a critical condition for assigning motor images and observed actions the status of covert and simulated actions is that they should activate brain areas known to be devoted to the preparation of actions. Experiments using neuro-imaging techniques show the existence of a common cortical and subcortical network activated during both motor imagery and action preparation (Jeannerod, 1999, 2001; Jeannerod and Frak, 1999). This network involves structures directly concerned with motor execution, such as primary motor cortex, dorsal and ventral premotor cortex, lateral cerebellum, and basal ganglia ; it also involves areas concerned with action planning, such as dorsolateral prefrontal cortex and posterior parietal cortex. A similar pattern of activation was found during action observation (Buccino *et al.*, 2001), including activation of the motor cortex itself (Fadiga *et al.*, 1995) (more on this below).

3.3 Simulation theory and self–identification: The problem of shared representations of action

The central monitoring theory and the simulation theory have some characteristics in common. For example, in one of its latest versions (Frith *et al.*, 2000), the central monitoring theory postulates that the desired movement is *simulated* by the internal model in order to predict the sensory feedback that would result from the movement if it were executed. The two theories, however, clearly depart from each other when the nature of the represented actions comes into consideration. As we stressed in the preceding section, recognition of action and self-identification are ultimately processes that involve the participation of several persons. Thus, the

point made by the simulation theory, but not by the central monitoring theory, is that actions of others should be represented to the same extent as one's own actions, and that because the two types of representations will be present in the same brain, they must be disentangled from each other. Self-identification mechanisms thus cannot operate from a solipsist point of view, they must consider the existence of other agents. In the central monitoring theory, self-identification can only be achieved through the interpretation of the signals related to an action: if these signals do not belong to the self, they must belong to someone else (a default attribution). In the simulation theory we are proposing here, specific mechanisms for attribution to another agent are postulated.

In principle, a theory that postulates that both actions of the self and actions of the other can be distinguished solely on the basis of their central representations should predict separate representations for these two types of actions. At the neural level, one should expect the existence of different networks devoted to agent-identification depending on whether the action originates from the self or not. One network would have the function of recognizing actions as belonging to the self. Another should correspond to the attribution of actions to another person. In section 3.2, we presented evidence that brain areas activated while representing self-produced actions (executed or not) and observing actions of other people were partly overlapping. For example, the motor brain appeared to be activated in both cases. Thus, according to the concept of 'shared representations', introduced by Daprati *et al.* (1997) and Georgieff and Jeannerod (1998), different mental states concerning actions (e.g. intending an action and observing it performed by another agent) appear to share the same neural representations.

To clarify this concept, let us briefly describe experimental results obtained from monkeys. Indeed, a dramatic illustration of what a shared representation can be is offered by the finding of mirror neurons (Di Pellegrino *et al.*, 1992; Rizzolatti *et al.*, 1995). Mirror neurons were identified in the monkey premotor cortex. They are activated in two conditions: first, they fire when the animal is involved in a specific motor action, like picking a piece of food with a precision grip; second, they also fire when the animal simply watches the same action performed by an external agent (another monkey or an experimenter). In other words, mirror neurons represent one particular type of action, irrespective of the agent who performs it. At this point, it could be suspected that the signal produced by these neurons, and exploited by other elements downstream in the information processing flow, would be the same for an action performed by the self and by another agent: the two modalities of that action (executed and observed) would thus completely share the same neural representation. The problem of agent-identification, however, is solved by the fact that other premotor neurons (the canonical neurons) and, presumably many other neuron populations as well, fire only when the monkey performs the action and not when it observes it performed by another agent. This is indeed another critical feature of the shared representations concept: they overlap only partially, and the non-overlapping part of a given representation can be the cue for attributing the action to the self or to the other.

The same mechanism operates in humans. Neuroimaging experiments where brain activity was compared during different types of simulated actions (e.g. intending actions and preparing for execution, imagining actions, observing actions performed by other people) revealed, first, that there exists a cortical network common to all conditions, to which the inferior parietal lobule (areas 39 and 40), the ventral premotor area (ventral area 6), and part of SMA contribute; and second, that motor representations for each individual condition are clearly specified by the activation of cortical zones which do not overlap between conditions (Decety *et al.*, 1994, 1997; Grafton *et al.*, 1996; Rizzolatti *et al.*, 1996a; Gérardin *et al.*, 1999; Ruby and Decety, 2001).

In this section, we argued that the motor system with its internal inverse and forward models computes representations of imagined and observed actions as well as representations of actions to be performed by the agent. A comparison of actual with predicted feedback cannot therefore, contrary to what the central monitoring hypothesis suggests, be the sole basis of self- vs. other-attribution. Since the distinction between self and other must also be made for purely covert actions, attribution mechanisms cannot operate from a solipsistic point of view; specific mechanisms for attribution to other agents should also exist. The evidence we have reviewed suggests that besides a common network activated in all conditions— i.e., whether the action is actually executed or merely imagined and whether it involves the self or another agent—they are also different and non–overlapping networks whose activation is specific to a given condition. Moreover, the existing evidence suggests that the posterior parietal cortex plays an important role in self- vs. other attribution.

4. Failure of self–recognition/attribution mechanisms in pathological states

In this section, we investigate the effects of pathological conditions as another potential source of information concerning the mechanisms of self-identification. Pathological conditions offer many examples of misattributions: a typical case is that of schizophrenia. Misattributions are not evenly distributed among schizophrenic patients. This type of disturbance is preferentially observed in one particular class of patients who present the so–called 'first-rank symptoms'. According to Schneider (1955), these symptoms refer to a state where patients interpret their own thoughts or actions as due to alien forces or to other people and feel being controlled or influenced by others. First-rank symptoms (or 'Schneiderian' symptoms, as they will be called here) reflect the disruption of a mechanism that normally generates consciousness of one's own actions and thoughts and makes possible their correct attribution to their author. Thus, a study of attribution behavior in schizophrenic patients would not only help us understand the factors responsible for misattribution in the patients, but it would also shed light on this critical function in normal life.

The pattern of misattributions in these patients is twofold. According to the French psychiatrist Pierre Janet (1937), patients may either attribute their own actions or thoughts to others rather than to themselves (under-attributions); or patients may attribute the actions or thoughts of others to themselves (over-attributions). A typical example of under-attributions is hallucinations. Hallucinating schizophrenic patients may show a tendency to project their own experience onto external events. Accordingly, they may misattribute their own intentions or actions to external agents. During auditory hallucinations, the patient will hear voices that are typically experienced as coming from a powerful external entity, but which in fact correspond to subvocal speech produced by the patient (see below). The voices are often comments where the patient is addressed in the third person and also include commands and directions for action (Chadwick and Birchwood, 1994). The patient may declare that he or she is being acted upon by an alien force, as if his or her thoughts or acts were controlled by an external agent. The so-called mimetic behavior observed at the acute stage of psychosis also relates to this category. Over-attributions, which Janet (1937) called 'excess of appropriation', correspond for the patient to the illusion that actions of others are in fact initiated or performed by him/her and that he/she is influencing other people (the clinical picture of megalomania). In this case, patients are convinced that their intentions or actions can affect external events, for example, that they can influence the thought and the actions of other people. Accordingly, they tend to misattribute the occurrence of external events to themselves. The consequence of this misinterpretation would be that external events are seen as resulting from their own intentions and actions.

4.1 Is there a failure of the action monitoring system in Schneiderian schizophrenic patients?

The most prevalent theory for explaining attribution difficulties in schizophrenic patients is that proposed by Frith and his colleagues. This theory, which was first formulated in the early 1990s (e.g. Frith, 1992) has been recently reformulated within the framework of the action monitoring mechanism described earlier in section 3.1 of this paper. The general idea is that the delusion of influence arises from a lack of awareness of the predicted limb position. The reasoning is the following: 'Under normal circumstances the awareness of initiating a movement must depend on the predicted limb position because awareness of initiating a movement precedes the actual movement and any feedback about actual limb position. The patient with delusions of control is aware of his goal, of his intention to move and of his movement having occurred, but he is not aware of having initiated the movement. It is as if the movement, although intended, has been initiated by some external force' (Blakemore *et al.*, 2002, p. 240). In terms of the action monitoring theory, the key problem for these patients would be that, due to their pathological condition, the endogenous signals related to action generation (the so-called efference copy) would be either absent or not properly used by the nervous system.

These contentions are supported by experimental evidence showing that Schneiderian patients fail to monitor the discrepancies between their intended and predicted limb positions in situations where they have to make rapid corrections. In these experiments, patients were shown moving visual targets that they had to reach by displacing a lever. No visual feedback about their movement was provided. The main result was that, if the target moved in an unexpected direction, the patients tended not to correct their error when they had started their movement in the direction where they expected the target to appear. By contrast, the same patients had no difficulty making the correction when they were provided visual feedback on the displacement of the lever (e.g. Frith and Done, 1989; Mlakar *et al.*, 1994). The problem with this type of experiment, however, is that it mostly stresses the automatic part of the action monitoring system, but tends to overlook the conscious part of this mechanism, that which makes it possible to give an explicit judgement on the action. In other words, the link between the experimentally proven inability to make corrections and the clinically observed misattributions of the action to its agent is not firmly established.

Before this question can be solved, however, one needs to address the issue of conscious awareness of the action in schizophrenic patients. With this aim in view, Fourneret *et al.* (2001) used a paradigm similar to that of Frith and Done (1989) for testing the ability not only to produce automatic corrections but also to make conscious judgements on the motor task. They used the apparatus previously described in section 2.1 for the experiment by Fourneret and Jeannerod (1998). The task consisted in drawing a line in the sagittal direction while a fixed bias of 15° to the right was introduced in the system. An opaque mask occluded the screen, such that the subjects could not see the line they drew except in the last one third of the trajectory to the target. Twenty trials were performed. Patients with and without Schneiderian symptoms, as well as normal control subjects were tested. It took approximately 10 trials for subjects from all three groups to learn to gradually compensate for the bias and to maintain the same strategy of moving the arm to the left throughout the task. This result indicates that the patients were able, in the absence of visual feedback, to initiate a hand movement in the proper direction to reach the target. This implies that they monitored their central command signals, modified them so as to achieve the task and maintained this modification across the trials. In addition, most of the patients in the Schneiderian group were found to be able to explicitly describe the strategy they had to use in order to reach the target. The Fourneret *et al.* (2001) experiment therefore demonstrates that conscious action monitoring can be found in Schneiderian patients, provided they are tested with the appropriate experimental design.

4.2 A specific attribution deficit in schizophrenic patients

The findings reported in the above section suggest that the explanation for the misattribution observed in Schneiderian patients should be looked for at a level of action consciousness directly related to their experience of agency, rather than at

the level of the conscious control of the action. To address this point, we will examine the results obtained with groups of schizophrenic patients in experiments specifically designed for testing self-recognition and attribution. The basic methodology for these experiments, already described in section 2.2., is to record agency judgements about movements that are shown to the subjects and that may correspond, or not, to their own movements. The first experiment of this type was made by Daprati *et al.* (1997): subjects were requested to make simple finger movements with their hidden right hand. At the same time, they were shown on a screen the movements of a right hand of an uncertain origin, that is, a hand that could equally likely belong to them or to someone else. Finally, the movements performed by that hand either could be the same as those the subjects had performed or could be different. In the condition with the maximum uncertainty about the ownership of the hand, Schneiderian patients made attribution errors in 80% of trials, i.e. they tended to misattribute the movements of the alien hand to themselves. The error rate in the same condition was 30% in control subjects and around 50% in non-Schneiderian patients. The results thus reveal the existence of impaired attribution of action in schizophrenia, especially in Schneiderian patients.

In the above experiment, however, only one hand was shown at a time to the patient: thus, the experimental situation privileged self-attribution responses, because it always referred to the patient as the putative agent of the action. It can be conjectured that, if another agent had been clearly involved in the experiment, the patient could have produced the opposite type of response, i.e. a response of underattribution. This possibility was tested in the experimental situation described in section 2.2 (van den Bos and Jeannerod, 2002). The patient's hand was shown along with another hand and the patient was requested to attribute one of the two hands to himself or to the other. Several conditions where the orientation of the hand was varied and where movements were performed, or not, by the hands, were used. In this situation, Schneiderian patients indeed tended to make more attribution errors, but their pattern of errors changed with respect to the previous experiment where only one hand was shown: they misattributed the hands more frequently to the other than to themselves. Normal controls and non-influenced patients made an equal number of errors in both directions (Farrer *et al.*, in preparation).

In these experiments, the errors in the determination of agency were strongly dependent on the availability of movements performed by the hands that were submitted to identification by the patients. This effect was already clearly present in normal subjects, as shown in section 2.2. In fact, the nature of Schneiderian symptoms suggests that errors in the determination of agency should not be limited to attribution of overt actions to their agents.

4.3 The depth of misattribution in schizophrenic patients

According to Schneider, the first-rank symptoms observed in schizophrenic patients represent false beliefs which lead to a feeling of depersonalisation by

impairing the distinction between the self and other people. Thus, misattribution should not be limited to actions involving overt behavior; it should also extend deeper into the mechanisms of action generation, i.e. to covert mental states related to possible actions intended by the self or other people. Following this suggestion, we are now going to examine the evidence for a possible misattribution of covert, non-executed, actions in such patients.

Among Schneiderian symptoms, verbal hallucinations represent a typical case of misattribution. As already mentioned at the beginning of section 4, it is known that auditory verbal hallucinations in schizophrenic patients are in fact related to the production of covert speech (inner speech) by the patient. Some hallucinated patients even show muscular activity in their laryngeal muscles (e.g. Gould, 1949; David, 1994). Thus, the patients perceive their inner speech as voices arising from an external source. McGuire *et al.* (1996) examined schizophrenic patients predisposed to hallucinations (whom they called 'hallucinators', a category that corresponds to our above Schneiderian groups), patients with no history of hallu-cinations (non-hallucinators) and normal controls, in tasks that entailed the moni-toring of inner speech. Patients had to either think simple sentences in their mind (inner speech) or to imagine themselves hearing the same sentences spoken by an alien voice (verbal imagery). The authors verified that no vocal utterance was produced by the patients. Brain activity was measured during these tasks, using PET. The main effect of the tasks was that, when compared to controls and non-hallucinators, patients predisposed to hallucinations showed a reduction of activity in their left middle temporal gyrus during the task of imagining sentences pro-nounced by another person. McGuire *et al.*'s conclusion was that this reduced activity in an area specifically devoted to the processing of spoken language could interfere with the patient's recognition of his own language as self-generated. The reasoning of the authors was that regions concerned with the generation of language (e.g. in the frontal lobe) normally inform speech perception areas of imminent language output, so that the subsequently perceived speech is recognized as self-generated. This process would be dysfunctional in hallucinators, hence their vulnerability to hallucinations. This result is complementary to another one obtained in Schneiderian patients while they were experiencing hallucinations. The results show that, when hallucinations occur (as signalled by the patients), brain metabolism increases in the primary auditory cortex (Heschl gyrus) on the left side (Dierks *et al.*, 1999) Thus, during verbal hallucinations, the auditory temporal areas remain active, which suggests that the nervous system in these patients behaves as if it were actually processing the speech of an external speaker, whereas self-generated inner speech is normally accompanied by a mechanism that decreases the responsiveness of primary auditory cortex.

These data speak more in favor of a defective simulation mechanism for attributing verbal thinking to the self or to another person, than of a defective action monitoring mechanism. Indeed, because in verbal hallucinations there is no vocal utterance, the hypothesis of a defective simulation mechanism seems more appropriate for explaining that the sense of agency must be functioning even in the

absence of comparison with external reafferences, and that actions can be monitored at the level of their representation, not only at the level of their execution. Accordingly, we would propose that the network which we normally use for attributing thoughts and intentions to their agent (be it ourselves or another person) is damaged in certain types of schizophrenic patients and therefore does not allow them to properly simulate the covert operations needed for attribution. This hypothesis would be consistent with the fact that, in clinical practice, the Schneiderian symptoms concern mostly non-executed actions.

5. Is there such a thing as immunity to error though misidentification?

In section 1, we argued that the problem of self-other discrimination, hence of self-identification, could arise not just for the self and others as objects, as it does in bodily self-identification, but also for the self and others as subjects—more specifically as agents—if the following conditions obtained:

(1)   The intentions of others can be directly perceived;
(2)   Our primary awareness of intentions, either ours or those of others, is as naked intentions;
(3)   The channels that provide the collateral information needed to determine who the author of the intention is are not entirely reliable and cannot be a source of identification-free first-person knowledge.

In sections 3 and 4, we have presented material we take as relevant to determining whether these conditions hold. It is now time for us to spell out more precisely how we take the empirical results we have reviewed to bear on these three issues. We will examine each one in turn, but before we do so we want to say something about our use of the term 'intention'.

The words 'intention' and 'to intend' are ambiguous and are used by philosophers in a variety of ways. For instance, as pointed out by Velleman (1989), 'intention' can be used to refer to the immediate goal of an action or to the agent's ultimate motivating desire. A weightlifter lifting a heavy barbell may be described as intending to lift it (immediate intention) or as intending to prove that he is the strongest man in his country (ultimate motivating desire). Some philosophers (e.g. Davidson, 1980) characterize intentions in an essentially backward looking way and will typically describe an intention as the conclusion of a practical deliberation whose premises exhibit the agent's reasons for performing an action of a certain type. In contrast, other philosophers tend to describe intentions in a more forward-looking fashion, in terms of its role in the production, control and guidance of the corresponding action. Another important distinction is between prior intentions and intentions–in–action (Searle, 1983). Prior intentions are conceived as states temporally separate from the action itself and as specifying the action one is committing oneself to in general, descriptive terms.

Intentions-in-action are contemporaneous with the actual (or simulated) action, are thus directed at a particular action token and specify it not in descriptive terms but in indexical and sensory-motor terms. The notion of intention we have in mind is of the immediate, forward-looking, intention-in-action variety and may be said to correspond to a representation of an action understood as a pragmatic representation of a goal or, more specifically, to the consciously accessible part of such a representation. Claims (1)–(3) should be read accordingly.

5.1 Can the intentions of others be directly perceived?

When intentions are understood in this way, the question whether we can directly perceive the intentions of others may be rephrased as the question whether when we observe someone else acting we can directly form a representation isomorphic to the representation the observed agent uses to guide and control his action. When one observes a bit of behavior, can one immediately see it as intentional behavior and can one identify the goal this behavior is directed at? There are at least two reasons why one may object to the idea that intentions can be directly perceived. One has its source in an adherence to the Cartesian picture of the relation between body and mind. According to this picture, human bodies are conceived as merely material objects and their behavior as being intrinsically non-intentional. What confers intentional properties on behavior are its inner mental accompaniments and causes. In other words, nothing intrinsic distinguishes a mere bodily happening from a piece of intentional behavior; the difference is one of causal antecedents. Since internal mental causes can't be directly observed they must be inferred, thus it doesn't make sense to say that when we observe an action, we can directly perceive the intention of the agent. To answer this objection, one must show, contra the Cartesian picture, that intentional bodily behavior is not intrinsically indistinguishable from non-intentional bodily behavior. In section 3.2, we have reviewed evidence that intentional behavior has indeed distinctive observable properties and that we are highly sensitive to these properties.

   The second objection to the idea that intentions can be directly perceived has its source in an acceptance, implicit or explicit, of what Susan Hurley (1998) has called the Classical Sandwich Model, the view that perception and action are separate and peripheral, one an input system and the other an output system, whereas cognition is central and separate from processing of perception and action.[11] On this view, perception and action systems are conceived as separate modules, each with their proprietary codes and processing rules. The principles according to which visual coding and motor coding group token visual patterns and motor response patterns respectively into types are completely different. Their respective classifications, dimensions of similarity and variations are thus disparate and the links between them arbitrary. One consequence of this view is that the extraction of the significance of visual patterns for action requires a complex translation

---

[11]   See also Brewer (1993) for a critique of this classical model.

between different types of contents. One further consequence is that when we are visually presented with an action performed by some other agent, the way this visual pattern will be typed by our perceptual system will bear no direct relation to the motor type the seen behavior is a token of. Thus, adherence to the Classical Sandwich Model implies that intentions cannot be directly perceived.

Given the enormous amount of evidence gathered in the last two decades in favor of the existence of a 'vision for action' system whose processing of visual information is aimed at extracting its motor significance (Milner and Goodale, 1995; Jeannerod, 1997; Jacob and Jeannerod, 2003), the credentials of the Classical Sandwich Model appear extremely doubtful. Moreover, as the results described in section 3.2 show, there are good empirical reasons to think that the vision for action system is activated not just when the agent is preparing to act on the environment but also when he or she observes other agents acting. Contra the Classical Sandwich Model then, perceptual and motor categories are not heterogeneous categories bearing only arbitrary relations to one another. Perception and action are closely integrated and when we visually perceive actions, we seem to be immediately sensitive to the distinctive properties of intentional behavior.

These data undermine one form of skepticism about our knowledge of other minds, skepticism that stems from the idea that this knowledge can only be inferential. The need for inferential mediation is taken to be the consequence of the existence of a gap. In the Cartesian picture, the gap is an ontological as well as a logical gap between mental states and physical behavior. For a proponent of the Classical Sandwich Model, it is a gap between perceptual and motor categories. What we have tried to argue so far is that in the case of intentions and intentional behavior at least, neither the existence of a physical–mental gap, nor the existence of a perceptual–motor gap should be accepted. As a result, we have no reason to maintain that intentions are undetectable by direct experience and must be inferred on the basis of behavior.

5.2 Naked intentions

If each time we were non-inferentially aware of an intention, this intention was necessarily ours, the problem of self-other discrimination wouldn't arise. If, however, we can be aware of both our intentions and those of others in the same way, namely as unattributed or 'naked' intentions, the problem of self-other discrimination does indeed arise. One may conceive of the content of an intention either as an impersonal representation encoding two elements, <action, goal>, or as a personal representation encoding three elements, <agent, action, goal>. There are in turn two possible ways for an intention to be naked. It could be the case either that intentions—or rather the specific variety of intentions we are concerned with here—are impersonal representations or that, although their form is <agent, action, goal>, the agent parameter can be left unspecified. These two possibilities are not mutually exclusive, since an intention that at one level takes the form <action, goal>, could at another level be redescribed as <x, action, goal>, where

x stands for an as yet unspecified agent. We will not attempt here to adjudicate between these two possibilities. Rather, we want to point out that the existence of shared representations of action is evidence for the existence of naked intentions, whatever their exact form may be. As we saw in section 3.3, the signal produced by mirror neurons appears to be the same for an action performed by the self and by another agent: the two modalities of that action (executed and observed) seem to completely share the same neural representation. Moreover, mirror neurons do not constitute the only instance of shared representations. As we have seen, brain imaging studies show that the cortical network common to intending actions and preparing for execution, imagining actions, and observing actions performed by other people is quite extended. Our contention is that this cortical network provides the basis for the conscious experience of goal-directedness—the primary awareness of intentions—but does not by itself provide us with a conscious experience of self- or other-agency. This latter experience has its basis in the activation of cortical zones that do not overlap between conditions.

5.3 Who?

We can be aware of an intention, without by the same token being aware of whose intention it is. Given the existence of shared representations, something more than the sole awareness of a naked intention is needed to determine its author. Ultimately, whether the problem of self-other discrimination is a problem worth worrying about depends on what further information is needed to attribute authorship of the intention to oneself or to someone else, and on the availability and reliability of that information. When the naked intention one is aware of yields an overt action, the extra information needed to establish authorship may be found in the outside world. The question 'Is this intention mine?' would then be answered by answering the question: 'Is this my body performing the corresponding action?' However, as we saw in section 2, this way of proceeding is less than fully reliable even for normal subjects and, as we saw in section 4, it can become seriously unreliable for schizophrenic patients and in particular Schneiderian patients. In social contexts especially, we may easily be fooled and since these are also the contexts where we are most likely to become aware of naked intentions through observing others acting, self-other discrimination can certainly become a problem.

However, it should be possible to experience a sense of agency even in the absence of overt action. What about the attribution of representations of covert non-executed actions? If, in their case, misattributions were impossible, if the central signals used to determine the source of these representations were always available and perfectly reliable, even the misattributions described in the previous paragraph would in principle be avoidable. But this is not the case. As the hallucinations experienced by schizophrenic patients show, attribution mechanisms

based purely on the monitoring of central signals are no more immune to error than attribution mechanisms that rely on the comparison of central signals with sensory reafferences.

In a nutshell then, the bad news for philosophers is that self-identification is after all a problem. In the domain of action and intention at least, there is no such thing as immunity to error though misidentification, whether for the self as object (sense of ownership) or for the self as agent (sense of agency). The mechanisms involved in self- and other-attribution may be reasonably reliable in normal circumstances, but they are not infallible. But together with the bad news come some good news or, to be more cautious, some hope for good news. Immunity to error through misidentification in the self-attribution of mental states is one of the main roots of the purported asymmetry between our knowledge of our own mental states and our knowledge of the mental states of others, the other main culprit being the inferential character of our knowledge of other minds. This asymmetry is thought to create a semantic problem. How could our mental terms and concepts ever be univocal if the criteria we appeal to when applying these terms and concepts to ourselves are radically different from the criteria we use when applying them to others? If we are right, our knowledge of other minds need not always be inferential and self-attribution of mental states is not always immune to error through misidentification. At least some obstacles to developing a unified account of mental concepts have been removed.

Institut des Sciences Cognitives
France

Institut Jean Nicod, CNRS-EHESS-ENS,
France

References

Baldwin, T. 1995: Objectivity, causality, and agency. In J.L. Bermúdez, A. Marcel and N. Eilan (eds.), *The Body and the Self.* Cambridge, MA: MIT Press, pp. 107–125.

Beardsworth, T., and Buckner, T. 1981: The ability to recognize oneself from a video-recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society*, 18, 19–22.

Bertenthal, B.I., Profitt, D.R., and Cutting, J.E. 1984: Infant sensitivity to figural coherence in biomechanical motions. *Journal of Experimental Child Psychology*, 37, 213–230.

Binet, A.1886: *La Psychologie du Raisonnement. Recherches Expérimentales par l'Hypnotisme.* Paris, Alcan.

Bisiach, E. and Berti A. 1987: Dyschiria: an attempt at its systematic explanation. In M. Jeannerod (ed.), *Neurophysiological and Neuropsychological Aspects of Spatial Neglect.* Amsterdam: North-Holland.

Blakemore, S.J., Wolpert, D.M. and Frith, C.D. 2002: A framework for understanding abnormalities in the control and awareness of action. *Trends in Cognitive Science* 6, 6, 237–242.

Botvinick, M. and Cohen, J., 1998: Rubber hands 'feel' touch that eyes see. *Nature*, 391, 756.

Brewer, B. 1993: The integration of spatial vision and action. In N. Eilan, R. McCarthy and B. Brewer (eds), *Spatial Representation*. Oxford: Blackwell, pp. 294–316.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G. and Freund, H.-J. 2001: Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13, 400–404.

Cassam, Q. 1997: *Self and World*. Oxford: Clarendon Press.

Chadwick, P. and Birchwood, M. 1994: The omnipotence of voices. A cognitive approach to auditory hallucinations. *British Journal of Psychiatry*, 164, 190–201.

Cole, J., and Paillard, J. 1995: Living without touch and information about body position and movement. Studies on deafferented subjects. In J. L. Bermúdez, A. Marcel and N. Eilan (eds), *The Body and the Self.* Cambridge, MA: MIT Press, pp. 245–266.

Cutting, J.E. and Koslowski, L.T. 1977: Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353–356.

Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J. and Jeannerod, M. 1997: Looking for the agent. An investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, 65, 71–86.

David, A.S. 1994: The neuropsychological origin of auditory hallucinations. In A.S. David and J.C. Cutting (eds), *The Neuropsychology of Schizophrenia*. Lawrence Erlbaum, Hove, pp. 269–313.

Davidson, D. 1980: *Essays on Actions and Events*. Oxford: Oxford University Press.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. and Fazio, F. 1997: Brain activity during observation of actions: influence of action content and subject's strategy. *Brain*, 120, 1763–1777.

Decety, J., Jeannerod, M., and Prablanc, C. 1989: The timing of mentally represented actions. *Behavioural Brain Research*, 34, 35–42.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., and Fazio, F. 1994: Mapping motor representations with PET. *Nature*, 371, 600–602.

Di Pellegrino G., Fadiga L., Fogassi L., Gallese V., and Rizzolatti G. 1992: Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91, 176–180.

Dierks, T., Linden, D.E.J., Jandl, M., Formisano, E., Goebel, R., Lanferman, H. and Singer, W. 1999: Activation of the Heschl's gyrus during auditory hallucinations. *Neuron*, 22, 615–621.

Fadiga, L., Fogassi, L., Pavesi, G. and Rizzolatti, G. 1995: Motor facilitation during action observation. A magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608–2611.

Farné, A., Pavani, F., Meneghello, F. and Ladavas, E., 2000: Left tactile extinction following visual stimulation of a rubber hand. *Brain*, 123, 2350–2360.

Farrer, C., Franck, N., Georgieff, N., Frith, C.D., Decety, J. and Jeannerod, M. 2003: Modulating the experience of agency: a PET study. *Neuroimage*, 18, 324–333.

Farrer, C., Franck, N., Georgieff, N., Tiberghien, G., Marie-Cardine, M., Daléry, J., D'Amato, T. and Jeannerod, M.: Confusing the self and the other. Impaired attribution of actions in patients with schizophrenia. (In preparation).

Fourneret, P. and Jeannerod, M. 1998: Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia*, 36, 1133–1140.

Fourneret, P., Franck, N., Slachewsky, A. and Jeannerod, M. 2001: Self-monitoring in schizophrenia revisited. *NeuroReport*, 12, 1203–1208.

Frak, V.G., Paulignan, Y. and Jeannerod, M. 2001: Orientation of the opposition axis in mentally simulated grasping. *Experimental Brain Research*, 136, 120–127.

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J. D'Amato, T. and Jeannerod, M. 2001: Defective recognition of one's own actions in schizophrenic patients. *American Journal of Psychiatry*, 158, 454–459.

Frith C.D. and Done D.J. 1989: Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine*, 19, 359–363.

Frith C.D., Blakemore S.J. and Wolpert D. M. 2000: Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society London B*, 355, 1771–1788.

Frith, C.D. 1992: *The Cognitive Neuropsychology of Schizophrenia*. Hove: Lawrence Erlbaum Associates.

Gallagher, S. 1986: Body image and body schema: A conceptual clarification. *Journal of Mind and Behavior*, 7, 541–554.

Gallese V. and Goldman, A. 1998: Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Science*, 2, 493–501.

Georgieff, N. and Jeannerod, M. 1998: Beyond consciousness of external reality. A 'Who' system for consciousness of action and self-consciousness. *Consciousness and Cognition*, 7, 465–477.

Gérardin, E., Sirigu, A., Lehéricy, S., Poline, J-B., Gaymard, B., Marsault, C., Agid, Y. and Le Bihan, D. 2000: Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex*, 10, 1093–1104.

Goldman, A.I. 1995: In defense of the simulation theory. In M. Davies and T. Stone (eds), *Folk-Psychology*. Oxford: Blackwell, pp. 191–206.

Goldman, A.I. 2002: Simulation theory and mental concepts. In J. Dokic and J. Proust (eds). *Simulation and Knowledge of Action*. Amsterdam: John Benjamins, pp. 1–20.

Gopnik, A. and Meltzoff, N. 1997: *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.

Gordon, R.M. 1995a: Folk psychology as simulation. In M. Davies and T. Stone (eds), *Folk-Psychology*. Oxford: Blackwell, pp. 60–73.

Gordon, R. M. 1995b: Simulation without introspection or inference from me to you. In M. Davies and T. Stone (eds), *Mental Simulation*. Oxford: Blackwell, pp. 53–67.

Gordon, R.M. 1996: 'Radical' simulationism. In P. Carruthers and P. K. Smith (eds), *Theories of Theories of Mind*. Cambridge: Cambridge University Press, pp. 11–21.

Gould, L.N. 1949: Auditory hallucinations in subvocal speech: objective study in a case of schizophrenia. *Journal of Nervous and Mental Diseases*, 109, 418–427.

Grafton, S.T., Arbib, M.A., Fadiga, L. and Rizzolatti, G. 1996: Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research*, 112, 103–111.

Harris, C. S., 1965: Perceptual adaptation to inverted, reversed and displaced vision. *Psychological Review* 72, 419–444.

Harris, P. 1995: From simulation to folk psychology: The case for development. In M. Davies and T. Stone (eds), *Folk-Psychology*. Oxford: Blackwell, pp. 207–231.

Heal, J. 1994: Simulation vs. theory theory: What is at issue? In C. Peacocke (ed.), *Objectivity, Simulation and the Objectivity of Consciousness*. Oxford: Oxford University Press, pp. 129–144.

Heal, J. 1995: How to think about thinking. In M. Davis and T. Stone (eds), *Mental Simulation*. Oxford: Blackwell, pp. 33–52.

Heal, J. 1996: Simulation, theory and content. In P. Carruthers and P. Smith (eds), *Theories of Theories of Mind*. Cambridge: Cambridge University Press, pp. 75–89.

Hurley, S. 1998: *Consciousness in Action*, Cambridge, Mass.: Harvard University Press.

Jacob, P. and Jeannerod, M. 2003: *Ways of seeing. The scope and limits of visual cognition*. Oxford: Oxford University Press.

James, W. 1950: *The Principles of Psychology*, 2 vol. New York: Dover.

Janet, P. 1937: Les troubles de la personnalité sociale. *Annales Médico-Psychologiques*, tome II, 149–200.

Jeannerod, M. and Frak, V.G. 1999: Mental simulation of action in human subjects. *Current Opinions in Neurobiology*, 9, 735–739.

Jeannerod, M. 1994: The representing brain. Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–245.

Jeannerod, M. 1997: *The Cognitive Neuroscience of Action*. Oxford, Blackwell.

Jeannerod, M. 1999: To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology*, 52A, 1–29.

Jeannerod, M. 2001: Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14, S103–S109.

Jeannerod, M. 2003: Consciousness of action and self-consciousness. A cognitive neuroscience approach. In J. Roessler and N. Eilan (eds), *Agency and Self-awareness. Issues in Philosophy and Psychology*. Oxford: Oxford University Press, pp. 128–149.

Johansson, G. 1973: Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201–211.

Kandel, S., Orliaguet, J.P., and Viviani, P. 2000: Perceptual anticipation in handwriting: the role of implicit motor competence. *Perception and Psychophysics*, 62, 4, 706–716.

Koslowski, L.T. and Cutting J.E. 1977: Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, 21, 575–580.

Leslie, A.M. 1994: Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211–38.

Lipps, T. 1903: *Aesthetik: Psychologie des Schönen und der Kunst*. Hamburg: Voss.

McDowell, J. 1982: Criteria, defeasibility and knowledge. *Proceedings of the British Academy*, 68, 455–79.

McGuire, P.K., Silbersweig, D.A., Murray, R.M., David, A.S., Frackowiak, R.S.J. and Frith, C.D. 1996: Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine*, 26, 29–38.

Milner, A.D. and Goodale, M.A. 1995: *The Visual Brain in Action*. Oxford: Oxford University Press.

Mlakar, J., Jensterle, J., and Frith, C.D. 1994: Central monitoring deficiency and schizophrenic symptoms. *Psychological Medicine*, 24, 557–564.

Nielsen, T.I., 1963: Volition: a new experimental approach. *Scandinavian Journal of Psychology*, 4, 225–230.

Orliaguet, J.P., Kandel, S., and Böe, L.J. 1997: Visual perception of motor anticipation in cursive hand writing: influence of spatial and movement information on the prediction of forthcoming letters. *Perception*, 26, 7, 905–912.

Pacherie, E. 1997: Du problème de Molyneux au problème de Bach-y-Rita. In J. Proust (ed.), *Perception et Intermodalité, Approches actuelles du Problème de Molyneux*. Paris, P.U.F, 1997, pp. 255–293.

Paillard, J. 1982: Le corps et ses langages d'espace: Nouvelles contributions psycho-physiologiques à l'étude du schéma corporel. In E. Jeddi (ed.), *Le Corps en Psychiatrie*, Paris: Masson, pp. 53–69.

Paillard, J. 1999: Body Schema and Body Image. A double dissociation in deafferented patient. In G.N. Gantchev, S. Mori and J. Massion (eds), *Motor Control. Today and Tomorrow*. Sofia: Academic Publishing House 'Prof. M. Drinov', pp. 197–214.

Parsons, L.M. 1994: Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology. Human Perception and Performance*, 20, 709–730.

Poincaré, H., 1907: *La science et l'hypothèse*. Paris: Flammarion.

Ramachandran, V.S. and Rogers-Ramachandran, D. 1996: Synaesthesia in phantom limbs induced with mirrors. *Proceedings of the Royal Society London*, B263, 377–386.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L., 1995: Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131–141.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D. and Fazio, F. 1996: Localization of grasp representations in humans by PET. 1. Observation versus execution. *Experimental Brain Research*, 111, 246–252.

Ruby, P. and Decety, J. 2001: Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neurosciences*, 4, 546–550.

Runeson, S. and Frykholm, G. 1983: Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender, recognition, and deceptive intention. *Journal of Experimental Psychology – General*, 112, 585–615.

Russell, J. 1995: At two with nature: Agency and the development of dualism. In J.L. Bermúdez, A. Marcel and N. Eilan (eds.), *The body and the self*. Cambridge, MA: MIT Press, pp. 127–151.

Schneider, K. 1955: *Klinische Psychopathologie*. Stuttgart: Thieme Verlag.

Searle, J. 1983: *Intentionality*. Cambridge: Cambridge University Press.

Shoemaker, S. 1996: *The First-person Perspective and Other Essays*. Cambridge: Cambridge University Press.

Sirigu, A., Duhamel, J.-R., Cohen, L., Pillon, B., Dubois, B. and Agid, Y. 1996: The mental representation of hand movements after parietal cortex damage. *Science*, 273, 1564–1568.

Slachewsky, A., Pillon, B., Fourneret, P. Pradat-Diehl, Jeannerod, M. and Dubois, B. 2001: Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *Journal of Cognitive Neuroscience*, 13, 332–340.

Sperry, R.W. 1950: Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43, 482–489.

Stich, S.P. and Nichols, S. 1995: Folk psychology: Simulation or tacit theory. In M. Davies and T. Stone (eds), *Folk-Psychology*. Oxford: Blackwell, pp. 123–158.

Strawson, P.F. 1959: *Individuals*. London: Routledge.

Strawson, P.F. 1966: *The Bounds of Sense*. London: Routledge.

van den Bos, E. and Jeannerod, M. 2002: Sense of body and sense of action both contribute to self recognition. *Cognition*, 85, 177–187.

Velleman, J.D. 1989: *Practical Reflection*. Princeton: Princeton University Press.

von Holst, E. and Mittelstaedt, H. 1950: Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften*, 37, 464–476.

Wegner, D. 2002: *The Illusion of Conscious Will*. Cambridge, Mass: MIT Press.

Wittgenstein, L. 1953: *Philosophical Investigations*, transl. G.E.M. Anscombe. Oxford: Blackwell.

Wittgenstein, L. 1958: *The Blue and Brown Books*. Oxford: Blackwell.

Wolpert, D.M., Ghahramani, Z. and Jordan, M. I. 1995: An internal model for sensorimotor integration. *Science*, 269, 1880–1882.